*Soil–Vegetation–Atmosphere Transfer Schemes and Large-Scale Hydrological Models* (Proceedings of a symposium held during the Sixth IAHS Scientific Assembly at Maastricht, The Netherlands, July 2001). IAHS Publ. no. 270, 2001.

# Improved conditioning of SVAT models with observations of infrared surface temperatures

**M. F. McCABE, S. W. FRANKS & J. D. KALMA**

*Department of Civil, Surveying and Environmental Engineering, University of Newcastle, Callaghan 2308, New South Wales, Australia*

e-mail: cemfm@civeng.newcastle.edu.au

**Abstract** This paper reports on the use of an observed record of infrared surface temperature in conditioning SVAT model predictions of evapotranspiration. An investigation into whether such a time series can be employed to provide information on the dynamics of the land surface is presented. It is shown that calibration of modelled fluxes to observed responses of latent and sensible heat provides some constraint on model predictions. Significantly, calibrating the model to observations of the surface temperature showed constraint in estimates of cumulative evapotranspiration and also in the reproduction of the observed latent heat record. This result demonstrates that measurements of infrared temperature have some potential in improving the ability of land surface models to characterize surface heat fluxes. Using the infrared surface temperature records required some modification and the concept of a temperature difference approach is examined.

**Key words** evapotranspiration; surface temperature; calibration; SVAT model; land surface flux

## INTRODUCTION

The modelling of energy fluxes at the land surface forms a critical component of many studies. The physical processes concerned have been investigated through the use of many different land surface schemes. The multitude of approaches that can be employed to provide land surface representations range from multi-layer physically-based models, providing detailed microclimate profiles, to semi-empirical and simple bucket type models. While the underlying physics behind these schemes is generally the same, the degree of complexity within them varies considerably.

Soil–vegetation–atmosphere transfer (SVAT) models provide one approach to describing energy flux interactions with the surface. The abundance of SVAT models in the literature is matched by their varying degrees of complexity, often reflected in the number of parameters required to run the model and the degree of empiricism used. While the structure and conceptualizations of these models may be significantly different, all are prone to uncertainties in predictions and parameterizations. Since parameter uncertainty is inevitable, it is expected that complex model structures are more susceptible to error propagation though their parameterizations.

One technique to address the inherent uncertainty in model predictions is to calibrate the model against measured fluxes. In order to constrain feasible parameterizations and model predictions, additional information is required. This investigation focuses on single objective calibration comparing model responses to observed records of latent and sensible heat fluxes. Additionally, model reproductions are assessed through conditioning to a temporal record of infrared surface temperatures.

It will be shown that conditioning of modelled aerodynamic temperatures to absolute values of infrared surface temperatures does not provide correct representation of the land surface behaviour. However, comparison with the diurnal rise in temperatures, termed the temperature difference, shows appreciable insight into flux behaviour, offering potential for parameter constraint and improved predictions.

## OBSERVATION PROGRAMME AND METHODOLOGY

### Description of data set

The First International Satellite Land-Surface Climatology Project (ISLSCP) Field Experiment (FIFE), was conducted during 1987 over a 15 × 15 km region near Manhattan, Kansas, USA. A comprehensive description of this project and site descriptions can be found in Sellers *et al.* (1992). The data set used in this report spans the period 6–21 August 1987, during an intensive field campaign (IFC-3). Meteorological data were extracted from Station ID-5 (2123-SAM), containing measurements of dry and wet-bulb temperatures, wind speed and direction, net radiation, rainfall and a measure of the radiative surface temperature. This station was chosen as it provided a complete record for the period under investigation and was located near a number of flux measurement systems. Measurements of the latent and sensible heat flux were obtained from a Bowen Ratio system (1916-BRS) located nearby. Surface temperature measurements were obtained from an Everest Model 4000 mulitplexed infrared thermometer (IRT), of which several were installed at selected AMS sites throughout the FIFE study area.

### SVAT model description

The TOPUP-SVAT model (Beven & Quinn, 1994; Franks *et al.*, 1999) was developed in response to the trend towards ever more complex SVAT structures. The underlying rationale behind models of increased complexity is that improved process representation will yield parameters which are easier to measure or estimate. However, this is not necessarily the case. SVAT models require effective values for the various parameters at patch, regional or larger scales which cannot be easily estimated. The philosophy behind the TOPUP-SVAT is to move towards striking a balance between representing the key physical processes affecting land surface interactions, while doing so in a parametrically refined manner.

The TOPUP-SVAT model is conceptually similar to simple bucket type structures, with the exception that a dynamic water table is included. The model incorporates the effects of near surface stability conditions for the calculation of aerodynamic resistance and utilizes a series of equations describing the various atmospheric and surface controls on evaporation. The model structure is novel in that it includes the representation of water supply from upslope areas, allowing the parameterization of a locally varying water table which can maintain moisture for evapotranspiration during dry periods.

There are four pathways of evaporation derived from the moisture stores in the model, which include an interception storage and a root zone. Capillary rise from the water table and evapotranspiration from the water table when in the root zone are also modelled.

## Modelling methodology

As the initial conditions of the numerous variables within most dynamic systems cannot be known with any degree of certainty, techniques must be developed to account for these uncertainties. One approach that has been proposed to deal with this situation is the Generalised Likelihood Uncertainty Estimation (GLUE) methodology (Beven & Binley, 1992). The philosophy behind this approach is relatively straightforward and can be described as follows.

While some model parameters in land surface schemes can be measured directly, many serve as conceptual representations of a physical process and as such cannot by specified exactly (Gupta *et al.*, 1999). Also, many of these parameters do not behave in a linear fashion when moving across scales, and producing approximate ranges provides the only realistic technique available to quantify these variables. For those parameters that cannot be directly measured, broad ranges representative of the patch scale can be constructed, thus characterizing the relative uncertainty in parameter measurements. Specification of feasible ranges for each model parameter is an acknowledgement of the inherent uncertainty in land surface representations at a variety of scales.

Once ranges have been derived for each variable, parameter sets are constructed in order to run the model. This is accomplished using Monte Carlo simulation, which randomly samples the specified ranges and produces multiple unique parameter combinations. In this investigation 10 000 parameter sets were produced from within broadly defined ranges (see Franks & Beven, 1999). Each of these was run with 16 days of meteorological forcing data to produce 10 000 associated model outputs. It is reasonable to assume that within these multiple simulations occurs a number of parameter combinations that reflect the land surface observations. The concept that many different combinations of model parameters may reproduce calibration/validation data is described as the equifinality problem (Beven, 1993). The key question is how to distinguish those model outcomes that adequately reflect what is occurring at the land surface from those that do not.

## Specifying a likelihood function

Franks & Beven (1999) suggest that to reduce uncertainty one can either (a) improve the estimates or measurements of representative parameters across the area of interest, or (b) utilize more information in assessing the acceptability of model simulations. While many parameter combinations may be rejected through comparison to calibration data, many will remain as acceptable simulators of the system under investigation. Following the latter approach, records of measured sensible and latent heat have been used. Additionally, infrared surface temperatures measured from a mast-mounted

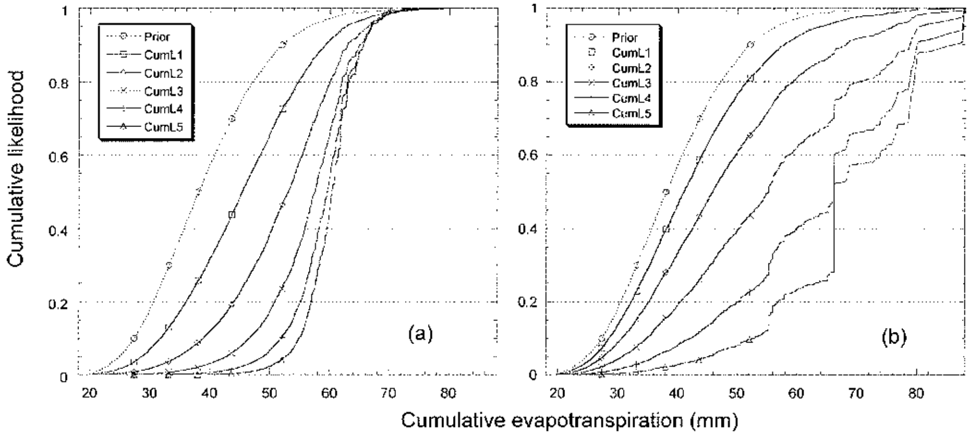

**Fig. 1** Cumulative likelihoods for conditioning to (a) latent heat and (b) temperature differences. The value of $N$ has been increased from 1 to 5, with the prior values representing equal likelihood value for all reproductions.

radiometer have been incorporated by comparing them to modelled aerodynamic surface temperatures. In order to compare observed and predicted records, a measure of the acceptability of model runs is required. This is termed the likelihood measure. The likelihood measure used here is a simple least squares error model comparing the observed and predicted responses for each parameter set, and is given by:

$$L(\theta|\underline{Y}) \propto \frac{SE_{\theta}^{-N}}{\sum_{10\,000}^{\theta=1} SE_{\theta}} \tag{1}$$

where $L(\theta|\underline{Y})$ is the likelihood estimate of parameter set $\theta$ for the set of forcing data $\underline{Y}$, and $SE_\theta$ is the squared error calculated for parameter set $\theta$. The power term $N$ is a subjectively chosen shaping parameter that accentuates the form of the response surface, as can be seen in Fig. 1. The squared error for each parameter set is calculated only when the input net radiation is greater than zero (during daylight hours). Uncertainty bounds specifying 5 and 95% provide a simple means of assessing model predictions by determining whether they envelop the observed data. In this study, the observed record was encompassed for all but the last few days, where predicted values just matched the observed response. Importantly, examination of all of the predicted data showed that simulated values were within the observed measures at all times.

## RESULTS

### Estimation of cumulative evapotranspiration

In order to examine any insight gained from conditioning model predictions to observed data, cumulative frequency plots were generated. These plots are useful in observing any insight into model predictions of surface fluxes or parameter values after calibration to measured variables. Cumulative evapotranspiration (*ET*) derived from the model output produced a range of predictions between 18 and 88 mm. The observed response had a cumulative *ET* of approximately 62 mm. While values of latent and sensible heat fluxes can be compared directly with model output, comparison of surface temperatures is hindered by differences in the physical nature of the measured variables. The aerodynamic surface temperature is not the same as an infrared surface temperature, although the difference between the two have been found to be typically less than 2°C over uniform canopy cover (Huband & Monteith, 1986). One technique put forward to account for this disparity is the idea of using a time rate of change in the temperature records (Diak & Whipple, 1995). Thus, by using the rate of change approach the need for absolute accuracy, which is unattainable, is reduced.

Modelled flux simulations were conditioned to observed records of latent and sensible heat and aerodynamic temperatures conditioned to measured infrared surface temperatures. In order to examine the benefit of utilizing infrared temperatures, both absolute values (using the entire record of surface temperatures) and a time rate of change (determined by the difference between a morning and afternoon temperature measurement) were examined as conditioning functions. Figure 1 shows the results of conditioning predictions of latent heat and the temperature difference to their observed responses.

Model outputs of latent and sensible heat once conditioned to their observed values produced a constraint in predictions around the observed value, indicated by the steepening of the lines in Fig. 1(a) for latent heat. This was expected and it provides an indication that the model is performing well. Conditioning to absolute values of surface temperature while producing some constraint, incorrectly identified the observed measurement of cumulative *ET* with predicted values almost 30% more than the observed measurement of 62 mm. On the other hand, using temperature differences produced predictions to within 5–10% of the observed value, as shown in Fig. 1(b). This illustrates the potential of utilizing a temperature difference approach and the inherent problems involved in using absolute values. The dotted line in Fig. 1 represents the prior likelihoods where each cumulative *ET* value is given equal likelihood weights, allowing an observation of the effects of conditioning on model predictions.

The value of the shaping parameter $N$ (equation (1)) was increased incrementally from 1 to 5 in Fig. 1. As $N$ increases, it moves towards the identification of an optimum set causing a constraint in the range of predicted cumulative *ET*. This allows the predicted ranges that best reflect the observed records to be identified. This is achieved by assigning those simulations that perform "better" with higher likelihoods than those that perform less well. While the specification of $N$ seems rather arbitrary, its use is in this paper is restricted to being a purely descriptive tool. Use of a high value of $N$ is really only justified if it is assumed that a model truly represents the process in question (Franks *et al.*, 1999). Otherwise, increasing $N$ will generally result in over conditioning towards a single optimum parameter set.

## Examining the temporal trend in flux predictions

While it has been shown that a temperature difference approach can correctly identify the *amount* of evapotranspiration, it is more important that it reproduces the pattern of surface fluxes. To examine this further, the optimum parameter set determined

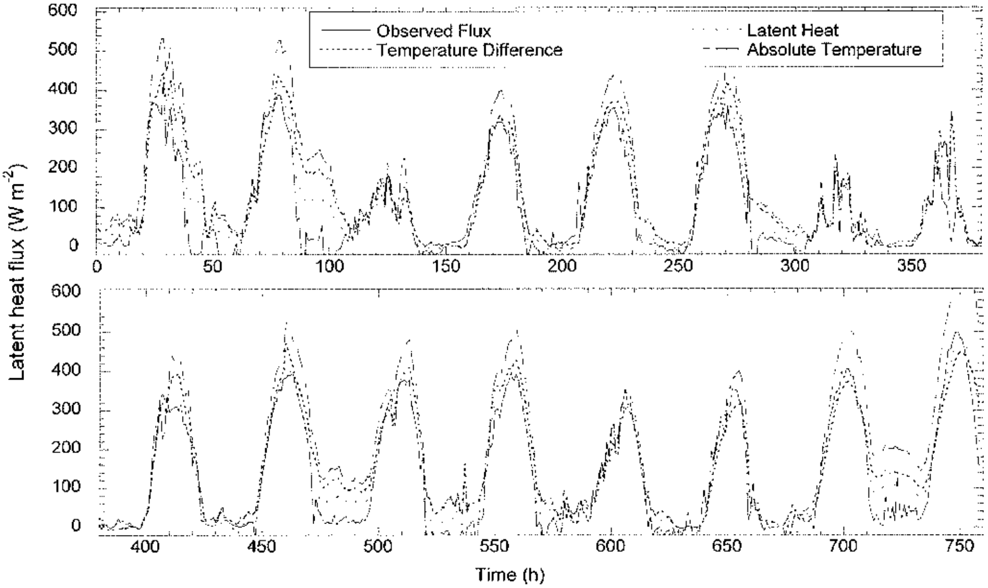

**Fig. 2** Comparison of conditioned predictions of latent heat, temperature differences and absolute temperature values compared to the observed record of latent heat flux.

through conditioning modelled aerodynamic temperature to both the observed infrared temperature difference and absolute values were identified. These simulations record of latent heat fluxes were then extracted and compared to that of the "best" predicted latent heat series. In doing this we are comparing like fluxes (latent heat) that have been conditioned on different records (infrared surface temperature and latent heat).

From Fig. 2, the record conditioned on infrared temperature differences is able to reproduce the observed pattern of latent heat flux as accurately as that conditioned on latent heat for all but the first two days of observations. Interestingly, it maintains this accuracy over a number of different climatic conditions, from a recently wetted soil (before time step 150), through to the drying period of the last three responses. Gupta *et al.* (1999) reported on single-criterion methods using the BATS model (Dickinson *et al.*, 1993), and found that model calibration to ground temperature did not reproduce observed fluxes for the ARM-CART and Tuscon, Arizona data sets. Importantly, it was concluded that single-criterion methods were found to be of limited value in constraining parameter ranges, particularly as they were not able to simulate the observations of other measured variables. This may be the result of conditioning to absolute values, which, as shown, do not accurately reproduce the latent heat record.

While the specification of an "optimum set" is not the aim of this calibration process, it does offer an opportunity to explore model performance. Future investigations into multi-conditioning model predictions to combinations of both measured fluxes and state variables (as opposed to the single conditioning explored here) should provide further insight into land surface behaviour. It is expected that this approach will also be effective in constraining parameter estimates, thus reducing model uncertainty and increasing efficiency.

## DISCUSSION

The characterization of the land surface within a modelling context is a complex problem. Issues related to the physical representation of system dynamics, the scale at which coupled processes are modelled, and the specification of appropriate parameter ranges, all complicate the effective simulation of land surface behaviour. The issue of model uncertainty is particularly pertinent, and much effort has been directed at developing techniques to assess the efficiency of model structure, and to reduce or identify ranges on model predictions.

The calibration of model predictions to observed records of latent and sensible heat illustrated that the model can effectively reproduce these fluxes. An examination of calibrating modelled aerodynamic surface temperature to the observed infrared surface temperature demonstrated its utility in providing insight into latent heat flux predictions. Conditioning predictions to absolute values of surface temperatures did not provide accurate insight into surface flux behaviour. However, by using the temperature difference approach, the disparity between these two variables was ameliorated, allowing a degree of land surface characterization to be achieved. An ability to assess surface flux predictions using infrared surface temperatures provides much potential in land surface modelling and also in generalized climate models, where the scale at which processes are modelled precludes the actual measurement of heat fluxes.

The ability of satellites to obtain information at regional scales provides an ideal platform from which to monitor land surface processes. Thermal infrared measurement can provide relatively accurate measures of the diurnal trend in land surface temperature (LST) at a range of spatial and temporal scales (McCabe *et al.*, 2000). Land surface temperatures derived from remote sensors have the potential to provide a valuable data source with which to initialize and calibrate a variety of modelling applications. Developing innovative procedures to make use of this information is required. The techniques employed in this study provide a means of assessing the information content of additional calibration sources and their ability to constrain model predictions, with the ultimate aim of providing insight into the modelling of surface fluxes and the production of more accurate, or less uncertain predictions.

## REFERENCES

Beven, K. J. (1993) Prophecy, reality and uncertainty in distributed hydrological modelling. *Adv. Wat. Resour.* **16**, 41–51.

Beven, K. J. & Binley, A. M. (1992) The future of distributed models: model calibration and uncertainty prediction. *Hydrol. Processes* **6**, 279–298.

Beven, K. J. & Quinn, P. F. (1994) Similarity and scale effects in the water balance of heterogeneous areas. Paper presented at AGMET Conf. on The Balance of Water—Present and Future (Dublin, Ireland).

Diak, G. R. & Whipple, M. S. (1995) Estimating surface sensible heat fluxes using surface temperatures measured from a geostationary satellite during FIFE 1989. Note, *J. Geophys. Res.* **100**(D12), 25,453–25,461.

Dickinson, R. E., Henderson-Sellers, A. & Kennedy, P. J. (1993) Biosphere Atmosphere Transfer Scheme (BATS) version 1e as coupled to the NCAR Community Climate Model. National Center for Atmospheric Research, Boulder, Colorado, USA.

Franks, S. W. & Beven, K. J. (1999) Conditioning a multiple patch SVAT model using uncertain time–space estimates of latent heat flux as inferred from remotely sensed data. *Wat. Resour. Res.* **35**, 2751–2761.

Franks, S. W., Beven, K. J. & Gash, J. H. C. (1999) Multi-objective conditioning of a simple SVAT model. *Hydrol. Earth System Sci.* **3**(4), 477–489.

Gupta, H. V., Bastidas, L. A., Sorooshian, S., Shuttleworth, W. J. & Yang, Z. L. (1999) Parameter estimation of a land surface scheme using multicriteria methods. *J. Geophys. Res.* **104**(D16), 19,491–19,503.

Huband, N. D. S. & Monteith, J. L. (1986) Radiative surface temperature and energy balance of a wheat canopy. I. Comparison of radiative and aerodynamic canopy temperature. *Boundary-Layer Met.* **36**, 1–17.

McCabe, M. F., Prata, A. J. & Kalma, J. D. (2000) A comparison of brightness temperatures derived from geostationary and polar orbiting satellites. *Tech. Report, CSIRO Division of Atmospheric Research, Aspendale, Victoria, Australia.*

Sellers, P. J., Hall, F. G., Asrar, G., Strebel, D. E. & Murphy, R. E. (1992) An overview of the First International Satellite and Land Surface Climatology Experiment (ISLSCP) Field Experiment (FIFE). *J. Geophys. Res.* **97**(D17), 18,345–18,372.